



AfIA

Association française
pour l'Intelligence Artificielle

Bulletin N° 107

Association française pour l'Intelligence Artificielle

AfIA

**Afia**Association française
pour l'Intelligence Artificielle

■ SISO : Système d'Information Spatialisé, Modélisation, Extraction et Diffusion des Données et Connaissances

TETIS UMR 9000 / SISO
AgroParisTech, CIRAD, CNRS, INRAE
<https://umr-tetis.fr>

Mathieu ROCHE
mathieu.roche@cirad.fr

Membres :

- Vincent BONNAL (CIRAD)
- Rémy DECOUPES (INRAE)
- Hugo DELEGLISE (doctorant)
- Vincent DOUZAL (INRAE)
- Sophie FORTUNO (CIRAD)
- Jean Eudes GBODJO (doctorant)
- Dino IENCO (INRAE)
- Roberto INTERDONATO (CIRAD)
- Rodrique KAFANDO (doctorant)
- Urcel KALENGA (doctorant)
- Martin LENTSCHAT (doctorant)
- Jérôme PASQUET (Univ. Paul Valéry)
- Mathieu ROCHE (CIRAD)
- Lucile SAUTOT (AgroParisTech)
- Maguelonne TEISSEIRE (INRAE)
- Sarah VALENTIN (doctorante)

Introduction

L'objectif de l'équipe SISO est de développer des méthodes de gestion de l'information permettant de répondre aux grands enjeux sociétaux liés à l'environnement et à l'agriculture, qu'il s'agisse de stocker, de gérer, de partager ou d'analyser de gros volumes de données. Les données et informations, décrites par des caractéristiques spatiales, temporelles et/ou thématiques, sont de surcroît hétérogènes ouvrant de nouvelles problématiques de recherche. Dans ce contexte, des contributions méthodologiques sont proposées et mises en place pour consolider la chaîne de l'information et les processus d'extraction de connaissances.

Activités

La multitude et la variété des données textuelles ainsi que l'émergence de nouvelles formes d'écriture rendent difficile l'extraction automatique d'information à partir de données textuelles sou-

vent hétérogènes et/ou de domaines spécialisés. Afin de relever ces défis, l'équipe SISO propose des approches originales de fouille de textes permettant l'identification automatique des informations spatio-temporelles et thématiques et leur mise en relation à partir de corpus mis à disposition auprès de la communauté scientifique sur des infrastructures de mutualisation et de partage de données numériques ([Dataverse](#), [Human-Num](#), [Ortolang](#)).

Extraction d'entités spatiales et thématiques

Une partie des travaux de l'équipe SISO consiste à proposer de nouvelles méthodes d'identification des entités spatiales (absolues et relatives) à partir de corpus peu standardisés [9] dans les domaines de l'agronomie [3] et de l'épidémiologie [1]. Ces informations spatiales peuvent être désambiguïsées par des méthodes d'apprentissage supervisé et d'apprentissage actif. Les travaux propres à l'identification d'informations thématiques reposent sur l'extraction de la terminologie (logiciel BioTex) via la proposition de nouvelles fonctions de rang [6], le labelling [8] et l'induction d'informations sémantiques [7]. Certaines méthodes mises en œuvre reposent sur la définition de nouveaux descripteurs linguistiques adossés à des méthodes d'apprentissage supervisé.

Mise en relation des entités

La mise en relation des différentes entités extraites est alors proposée. Dans ce cadre, des structures appelées STR (*Spatial Textual Representation*) permettent de représenter, de manière automatique, la configuration spatiale d'un document par des graphes dont les nœuds sont les entités spatiales désambiguïsées et les arcs les différentes relations spatiales (adjacence, inclusion, etc.) [4].



Les relations entre entités sont aussi modélisées et extraites sous forme de relations n-aires en combinant des méthodes de fouille de données (extraction de motifs et règles séquentiels) et d'analyse syntaxique [2]. Enfin, l'ensemble des entités (spatio-temporelles et thématiques) mises en relation constituent des événements pour des applications en veille épidémiologique réalisées dans un cadre pluridisciplinaire [1]. Ainsi, des logiciels de veille en épidémiologie animale (PADI-Web, Epid-News, EpidVis) accompagnés de nouvelles visualisations ont été produits [1, 5].

Références

- [1] Elena Arsevska, Sarah Valentin, Julien Rabatel, Jocelyn de Goer de Hervé, Sylvain Falala, Renaud Lancelot, and Mathieu Roche. Web monitoring of emerging animal infectious diseases integrated in the french animal health epidemic intelligence system. *PloS One*, 13(8), 2018.
- [2] Soumia Lilia Berrahou, Patrice Buche, Juliette Dibie, and Mathieu Roche. Xart : Discovery of correlated arguments of n-ary relations in text. *Expert Systems with Applications*, 73 :115 – 124, 2017.
- [3] Brett Drury and Mathieu Roche. A survey of the applications of text mining for agriculture. *Computers and Electronics in Agriculture*, 163 :104864, 2019.
- [4] Jacques Fize, Mathieu Roche, and Maguelonne Teisseire. Mapping heterogeneous textual data : A multidimensional approach based on spatiality and theme. In *Proc. Internet Science*, pages 310–317, 2019.
- [5] Rohan Goel, Sarah Valentin, Alexis Delaforge, Samiha Fadloun, Arnaud Sallaberry, Mathieu Roche, and Pascal Poncelet. Epidnews : Extracting, exploring and annotating news for monitoring animal diseases. *Journal of Computer Languages*, 2019.
- [6] Juan Antonio Lossio-Ventura, Clement Jonquet, Mathieu Roche, and Maguelonne Teisseire. Biomedical term extraction : overview and a new methodology. *Information Retrieval Journal*, 19(1) :59–99, 2016.
- [7] Juan Antonio Lossio-Ventura, Clement Jonquet, Mathieu Roche, and Maguelonne Teisseire. A novel framework for biomedical entity sense induction. *Journal of Biomedical Informatics*, 84 :31 – 41, 2018.
- [8] Julien Velcin, Antoine Gourru, Erwan Giry-Fouquet, Christophe Gravier, Mathieu Roche, and Pascal Poncelet. Readitopics : Make your topic models readable via labeling and browsing. In *Proc. of IJCAI*, pages 5874–5876, 2018.
- [9] Sarah Zenasni, Eric Kergosien, Mathieu Roche, and Maguelonne Teisseire. Spatial information extraction from short messages. *Expert Systems with Applications*, 95 :351 – 367, 2018.

**Afia**Association française
pour l'Intelligence Artificielle

CONSEIL D'ADMINISTRATION

Yves DEMAZEAU, *président*
Domitile LOURDEAUX, *vice-présidente*
Catherine FARON-ZUCKER, *trésorière*
Sandra BRINGAY, *secrétaire*
Dominique LONGIN, *rédacteur*
Emmanuel ADAM, *webmestre*

Membres :

Alain BERGER, Grégory BONNET, Élise BONZON,
Pierre FEILLET, Marie LEFÈVRE, Frédéric MARIS,
Engelbert Mephu NGUIFO, Gauthier PICARD, Cé-
line ROUVEIROL, Laurent SIMON, Olivier SIMONIN,
Charlotte TRUCHET

COMITÉ DE RÉDACTION

bulletins@afia.asso.fr

Emmanuel ADAM
Rédacteur
Grégory BONNET
Rédacteur en chef adjoint
redacteur-adjoint@afia.asso.fr
Claire LEFÈVRE
Rédactrice
Dominique LONGIN
Rédacteur en chef
redacteur@afia.asso.fr
Laurent SIMON
Rédacteur

LABORATOIRES ET SOCIÉTÉS ADHÉRANT COMME PERSONNES MORALES

.....
Ardans, Berger Levraut, CRIL, CRISTAL, Dassault Aviation, ENIB, EURODECISION, GRETTIA,
GREYC, Huawei, I3S, IBM, INRIA Sophia Antipolis Méditerranée, IRIT, ISAE-SUPAERO, Lab-STICC,
LAMSAD, LERIA, LGI2P, LHC, LIG, LIMICS, LIMSI, LIP6, LIPADE, LIRIS, LIRMM, LITIS, MaIAGE,
Naver Labs, Renault, Thales, Université Paris-Saclay, Veolia.

Pour contacter l'Afia

Président

Yves DEMAZEAU
L.I.G./C.N.R.S., Maison Jean Kuntzmann
110, avenue de la Chimie, B.P. 53
38041 Grenoble cedex 9
Tél. : +33 (0)4 76 51 46 43
president@afia.asso.fr

Serveur WEB

<http://www.afia.asso.fr>

Adhésions, liens avec les adhérents

Catherine FARON-ZUCKER
tresorier@afia.asso.fr

Calendrier de parution du Bulletin de l'Afia

	Hiver	Printemps	Été	Automne
Réception des contributions	15/12	15/03	15/06	15/09
Sortie	31/01	30/04	31/07	31/10